On Simple Estimation of the Parameters of the Weibull or Extreme-Value Distribution*

Max Engelhardt

Department of Mathematics
University of Missouri-Rolla
Rolla, Missouri 65401

Simple, closed form approximations for maximum likelihood estimates of the parameters of the Weibull or extreme-value distribution are discussed. A method for the exact computation of constants required to calculate the estimates is presented, and simpler approximate methods are also provided. Some inference procedures for the parameters are also discussed.

Key Words
Weibull Distribution
Extreme-Value Distribution
Censored Sampling
Estimation
Inference Procedures

1. Introduction and Summary

In a series of papers, Bain [3], and Engelhardt and Bain [7, 8] have developed simple, yet rather efficient unbiased estimators for the parameters of the extreme-value distribution having distribution function

$$F_Y(y) = 1 - \exp \left[ -\exp \left( \frac{y - u}{b} \right) \right],\quad -\infty < y < \infty, -\infty < u < \infty, b > 0.$$  

Bain [3] proposed an estimator \( \hat{b} \), based upon the smallest \( r \) out of \( n \) ordered observations, for the scale parameter \( b \). It was shown by Engelhardt and Bain [7] that \( \hat{b} \) has zero asymptotic efficiency when applied to a complete sample and a generalized estimator, with better asymptotic properties, was proposed. Engelhardt and Bain [8] also proposed an estimator \( \hat{u} \) for the location parameter \( u \). It was noted that \( \hat{u} \), as proposed for censored sampling, has zero asymptotic efficiency when applied to a complete sample and a modified form was provided for complete sampling.

Interest in estimators of this type is evident from other related papers. Mann and Fertig [14] independently and concurrently proposed the same estimator, from a censored sample, for the location parameter \( u \). It was also learned recently that estimators which coincide with \( \hat{u} \) and \( \hat{b} \), in the censored case, were studied by Abe [1], who had earlier used these estimators to estimate the expected maximum volume (during a period) of falling rocks on a mountain-side railway. This application was published (in Japanese) in [2].

A close agreement was noted in [8] between the maximum likelihood estimators (MLE’s) \( \hat{u} \) and \( \hat{b} \), and the respective estimators \( u^* = \hat{u} - \hat{b} \) \( \text{Cov}(\hat{u}, \hat{b})/(1 + \text{Var}(\hat{b})) \) and \( b^* = \hat{b}/(1 + \text{Var}(\hat{b})) \). The estimator \( b^* \) is also the simple approximation for the best linear invariant estimator (BLIE) of \( b \) which was discussed in [3], and \( u^* \) is a simple approximation for the BLIE of \( u \). It is well known and easily verified that the logarithm of a Weibull distributed random variable has an extreme-value distribution. Consequently, \( X = \exp(Y) \) is Weibull distributed with scale parameter \( \alpha = \exp(u) \) and shape parameter \( \beta = 1/b \). This suggests approximate BLIE’s \( \alpha^* = \exp(u^*) \) for \( \alpha \), \( \beta^* = 1/b^* \) for \( \beta \) and \( R^*(t) = \exp \left[ -\exp \left( \frac{\ln t - u^*}{b^*} \right) \right] \) for Weibull reliability \( R(t) \). Estimators having the same form as \( u^* \) and \( b^* \), in the censored case, are also discussed by Mann, Schaefer and Singpurwalla [15, p. 210] where the Monte Carlo method was used to obtain the moments required in calculating the estimates (see Table 5.11, [15, p. 244]).

An interesting comparison can be made between the approximate BLIE’s \( u^* \) and \( b^* \), the MLE’s \( \hat{u} \) and \( \hat{b} \), and the estimators \( \hat{u}' \) and \( \hat{b}' \) of Johns and Lieberman [10] which are in fact, asymptotic approximations to the BLIE’s. In regard to the latter estimators, D’Agostino [6] gave a simplified method for generating the necessary coefficients and tabulated mean squared errors (MSE’s) for sample sizes \( n = 3(1)10 \). The MSE’s of \( \hat{u} \) and \( \hat{b} \), based upon Monte Carlo simulation, were provided by Harter and Moore [9] for \( n = 10 \) and 20. A comparison of MSE’s is made in Table 1 for sample

---

* Research was sponsored in part by the Aerospace Research Laboratories, Air Force Systems Command, United States Air Force, Contract Number F33615-73-C-4143.

Received July 1974; revised February, 1975.
out of n ordered extreme-value observations. The generating functions, for the calculation of the variances and covariance of \( \hat{u} \) and \( \hat{b} \), and estimators of Johns and Lieberman \( \hat{u}' \) and \( \hat{b}' \) for sample size ten.

<table>
<thead>
<tr>
<th>( r/n )</th>
<th>( u^* )</th>
<th>( \hat{u} )</th>
<th>( u^* )</th>
<th>( \hat{u} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>.2</td>
<td>2.317</td>
<td>2.354</td>
<td>2.317</td>
<td>2.354</td>
</tr>
<tr>
<td>.3</td>
<td>.950</td>
<td>.951</td>
<td>.950</td>
<td>.951</td>
</tr>
<tr>
<td>.5</td>
<td>.498</td>
<td>.515</td>
<td>.498</td>
<td>.515</td>
</tr>
<tr>
<td>.6</td>
<td>.355</td>
<td>.335</td>
<td>.355</td>
<td>.335</td>
</tr>
<tr>
<td>.7</td>
<td>.212</td>
<td>.226</td>
<td>.212</td>
<td>.226</td>
</tr>
<tr>
<td>.8</td>
<td>.164</td>
<td>.166</td>
<td>.164</td>
<td>.166</td>
</tr>
<tr>
<td>.9</td>
<td>.138</td>
<td>.137</td>
<td>.138</td>
<td>.137</td>
</tr>
<tr>
<td>1.0</td>
<td>.125</td>
<td>.122</td>
<td>.125</td>
<td>.122</td>
</tr>
</tbody>
</table>

In section 4 an exact method, based upon moment d which minimize the MSE of \( d/3* \) are also obtained. Sizes \( n \leq 10 \) than the tables in [15], and yield values of substantially greater accuracy. In section 3 inference procedures for \( \hat{u} \) and \( \hat{b} \) are discussed, and factors \( a \) which make \( \hat{a}^* \) unbiased for \( \beta \) and \( d \) which minimize the \( \text{MSE of } \hat{d}^* \) are also obtained. In section 4 an exact method, based upon moment-generating functions, for the calculation of the variances and covariance of \( \hat{u} \) and \( \hat{b} \) is given.

2. Computation of the Estimates

Let \( y_1 < y_2 < \cdots < y_r \) denote the smallest \( r \) out of \( n \) ordered extreme-value observations. The estimator

\[
\hat{b} = \frac{\sum_{i=1}^{r} |y_i - y_r| / nk_{r,n}}{r - 1}
\]

was proposed in [3] where the unbiased constants \( k_{r,n} \) were tabulated for \( n = 5(5)20, 30, 60, \) and 100. Values of \( k_{r,n} \) based upon Monte Carlo simulation, are also tabulated in [15, p. 244] for \( n = 25(5)60 \).

An approximation which is nearly correct to four significant figures, is provided here. The following approximation is based upon a quadratic fit to exact values \( k_{r,10} \), \( k_{r,20} \), and the asymptotic value \( k_{r} \) with \( r/n \rightarrow p \) as \( n \rightarrow \infty \):

\[
k_{r,n} = A_0 + A_1(10/n) + A_2(10/n)^2.
\]

Since \( k_{r,n} \) is not defined for \( r = 1 \) and \( n = 10 \), a fit to \( k_{r,20}, k_{r,30} \), and \( k_{r} \) is given for the case \( r/n = .1 \). The coefficients \( A_0, A_1, \) and \( A_2 \) are provided in Table II for \( r/n = .1(1)9 \). There is rather close agreement with the exact values provided in [3]. In particular, the approximate values are correct to four significant figures except for the case \( n = 60 \) and \( r/n = .9 \) where it differs by only a single unit in the last place.

The estimator \( \hat{u} = y_r - c_r \hat{b} \) was proposed in [8], for censored samples, where the unbiased constants \( c_r \) are the expected values of standardized extreme-value order statistics. These are provided by White [16] for \( n = 1(1)50(5)100 \), and Mann [13] for \( n = 1(1)25 \). Simulated values are also tabulated in [15, p. 244] for \( n = 25(5)60 \). The following approximation is based upon a quadratic fit to exact values of \( c_r \) for \( n = 20, 30, \) and \( \infty \) when \( r/n = .1, n = 10, 20 \) and \( \infty \) when \( .1 < r/n \leq .9 \):

\[
c_{r,n} = B_0 + B_1(10/n) + B_2(10/n)^2.
\]

The coefficients \( B_0, B_1, \) and \( B_2 \) are provided in Table 2. Comparing with exact values in [16] for \( r/n < .9 \), every approximate value is either correct to four significant figures or it differs by a single unit in the last place. For \( r/n = .9 \), every value is either correct to three significant figures or it differs by a single unit in the last place.

Exact values of \( n \text{ Var (b/b)} \) for \( n = 10, 20 \), and asymptotically as \( n \rightarrow \infty \) with \( r/n \rightarrow p \) are provided in [7], and the quadratic fit approach was applied directly to \( n \text{ Var (b/b)} \) for \( r/n = .3(1)9 \). The approximation \( n \text{ Var (b/b)} \) is recommended for \( r/n < .3 \). By fitting, instead, the reciprocal values \( 1/n \text{ Var (b/b)} \), a better overall approximation is obtained, with the major improvement occurring for heavier censoring. The following approximation is based upon a quadratic fit to exact values of \( 1/n \text{ Var (b/b)} \) for \( n = 20, 30 \) and \( \infty \).
when \( r/n = .1 \), and \( n = 10, 20 \) and \( \infty \) when \( .1 < r/n \leq .9 \):

\[
1/n \ Var (\hat{b}/b) = C_0 + C_1(10/n) + C_2(10/n)^2.
\]

The coefficients \( C_0 \), \( C_1 \), and \( C_2 \) are provided in Table 2. Comparisons with exact values indicate that the remarks concerning the approximation of \( c_{rs} \) also apply here.

The quadratic fit approach, applied to \( n \ Cov (\hat{a}/b, \hat{b}/b) \), does not yield an approximation with accuracy comparable to those already given. However, \( n \ Cov (y_r/b, \hat{b}/b) \) can be approximated with comparable accuracy. Approximate values of \( c_{rs} \), \( n \ Var (\hat{b}/b) \), and \( n \ Cov (y_r/b, \hat{b}/b) \) can then be combined in the following formula:

\[
n \ Cov (\hat{a}/b, \hat{b}/b) = n \ Cov (y_r/b, \hat{b}/b) - c_{rs} n \ Var (\hat{b}/b).
\]

The following approximation is based upon a quadratic fit to exact values of \( n \ Cov (y_r/b, \hat{b}/b) \) for \( n = 20, 30 \) and \( \infty \) when \( r/n = .1 \) and \( n = 10, 20 \) and \( \infty \) when \( .1 < r/n \leq .9 \):

\[
n \ Cov (y_r/b, \hat{b}/b) = D_0 + D_1(10/n) + D_2(10/n)^2.
\]

The coefficients \( D_0 \), \( D_1 \), and \( D_2 \) are provided in Table 2.

In summary, the approximations for \( k_{rs} \), \( c_{rs} \), \( n \ Var (\hat{b}/b) \), and \( n \ Cov (\hat{a}/b, \hat{b}/b) \) yield values which are nearly correct to four significant figures for \( r/n < .9 \), and three significant figures for \( r/n = .9 \). Since the formulas are exact at \( 10/n = 0 \), the approximations will be correct asymptotically. The exact fits at \( 10/n = 1/2 \) and 1 provide close approximations for \( n \geq 10 \) and convenient tabulation for \( r/n = .1(1.9) \). Comparisons with exact values, based upon the results in section 4, reveal that the values of \( n \ Var (\hat{b}/b) \) and \( n \ Cov (\hat{a}/b, \hat{b}/b) \), using the quadratic approximations, are substantially more accurate than the Monte Carlo values tabulated in [15].

The method is illustrated by a numerical example based upon the computer generated sample of 40 ordered extreme-value observations with \( \alpha = 100 \) and \( \beta = 2 \) provided by Harter and Moore [9]. Suppose the largest 20 observations are censored. Then \( \hat{b} = [(19)(4.220) - 69.42]/[40(5.584)] = .4817 \), and \( \hat{u} = 4.220 - (-.4106)(.4817) = 4.418 \). Since \( n \ Var (\hat{b}/b) = 1.819 \) and \( n \ Cov (\hat{a}/b, \hat{b}/b) = 1.064 \), we have \( b^* = .4817/[1 + .0455] = .461 \), and \( u^* = 4.418 - (.461)(.0266) = 4.41 \), \( \beta^* = 1/.461 = 2.17 \), \( \alpha^* = \exp (.441) = 82.3 \), and \( R^*(32.46) = \exp [\exp [(3.48 - 4.41)/.461]] = .875 \). For comparison, the MLE's are \( \hat{a} = 4.43 \), \( \hat{b} = .478 \), \( \hat{\alpha} = 83.9 \), \( \hat{\beta} = 2.09 \), and \( \hat{R}(32.46) = .872 \).

A generalized estimator \( \hat{b} = \sum_{r=1}^{n} y_r/\ nk_{rs} \), where \( s \) is chosen to minimize the variance of \( \hat{b} \), was proposed in [7]. For the censored sample case \( \hat{b} \) reduces back to \( \hat{b} \) as proposed in [3] and the quadratic approximations are applicable. However, quadratic approximations are not feasible for the complete sample case. It can be seen from the small-sample results in [7] that the values of \( k_{rs} \), \( c_{rs} \), \( n \ Var (\hat{b}/b) \) fluctuate as \( n \) increases. This is due to the variability of the largest order statistic for smaller \( n \), and the integral jumps in \( n - s \) for larger values of \( n \). Similar remarks hold for the following estimator which was proposed in [8] for complete samples: \( \hat{b} = -\hat{y} + y/\ nk_{rs} \), where \( \hat{y} \) is the sample mean and \( y = .5772 \) is the Euler constant. The

| TABLE 2—Coefficients for the quadratic approximations of \( k_{rs} \), \( c_{rs} \), \( 1/n \ Var (\hat{b}/b) \), and \( n \ Cov (y_r/b, \hat{b}/b) \). |
|---|---|---|---|---|---|---|---|---|---|
| \( r/n \) | .1 | .2 | .3 | .4 | .5 | .6 | .7 | .8 | .9 |
| \( A_0 \) | .10265 | .21129 | .32723 | .45234 | .58937 | .74274 | .92026 | 1.1392 | 1.4456 |
| \( A_1 \) | -.10274 | -.10622 | -.11060 | -.11554 | -.12025 | -.12540 | -.13031 | -.13540 | -.14057 | -.14629 |
| \( A_2 \) | .00001 | .00004 | .00007 | .00010 | .00014 | .00018 | .00023 | .00028 | .00033 | .00038 |
| \( B_0 \) | -.2.2504 | -.4.9999 | -.7.0309 | -.9.7173 | -.12.6651 | -.15.6421 | -.18.5313 | -.21.4513 | -.24.3913 |
| \( B_1 \) | -.54549 | -.7.0340 | -.8.7289 | -.10.7131 | -.12.6621 | -.14.5991 | -.16.4991 | -.18.3871 | -.20.2321 |
| \( B_2 \) | -.0.7848 | -.0.8186 | -.0.9767 | -.0.9335 | -.0.9993 | -.0.9993 | -.0.9993 | -.0.9993 | -.0.9993 |
| \( C_0 \) | .10261 | .21059 | .32569 | .45055 | .57882 | .71666 | .87281 | 1.0.299 | 1.1697 |
| \( C_1 \) | -.10263 | -.10554 | -.10873 | -.11218 | -.11571 | -.11952 | -.12362 | -.12782 | -.13226 |
| \( C_2 \) | .00002 | .00004 | .00006 | .00008 | .00010 | .00012 | .00014 | .00016 | .00018 |
| \( D_0 \) | .25973 | .27113 | .28480 | .30160 | .32305 | .35188 | .39384 | .46402 | .62397 |
| \( D_1 \) | -.0.1259 | -.0.1436 | -.0.1681 | -.0.2026 | -.0.2537 | -.0.3265 | -.0.4587 | -.0.8394 | -.21.509 |
| \( D_2 \) | .00044 | .00046 | .00067 | .00102 | .00162 | .00280 | .00550 | .00383 | .05934 |

TECHNOMETRICS®, VOL. 17, NO. 3, AUGUST 1975
results provided in [7] can be used to obtain values of $s$, $k_{1,n}$, and $\text{Var} (\hat{b}_{i}/b)$. For example, if $n = 10$, then $s = 10, b_{10,10} = 1.567$ and $\text{Var} (\hat{b}_{i}/b) = 0.0795$, and if $n = 20$, then $s = 19, b_{19,20} = 1.520$ and $\text{Var} (\hat{b}_{i}/b) = 0.0402$. Exact covariances for $n = 10$ and 20 are $\text{Cov} (\hat{a}/b, \hat{b}/b) = -.0142$ and 0.0074 respectively. To calculate the approximate MLE’s for $n > 20$, in the complete sample case, it is convenient simply to use the asymptotic approximations $s = \lfloor 8.92 \ln n \rfloor + 1, k_{1,n} = 1.449, \text{Var} (\hat{b}_{i}/b) = .800/n$ and $\text{Cov} (\hat{a}/b, \hat{b}/b) = -.162/n.$

An estimator for $u$, in the complete sample case, is also proposed in [15, p. 211]. For small samples ($n \leq 15$) this estimator agrees with $\hat{u}$ as defined above. However, for larger $n$, $\hat{u}$ is more efficient. In particular, for $n = 20$ $\hat{u}$ has efficiency .93 (relative to the best linear unbiased estimator), compared to .89 for the estimator in [15]. The respective asymptotic efficiencies (relative to the Cramér–Rao lower bound) are .95 and .86.

The complete sample method will be illustrated with the data from [9]. Since $n > 20$, we use $s = \lfloor 8.92 \ln n \rfloor + 1, b_{n} = (31)(4.956) - 143.4 + 20.32)/[(40)(1.449)] = .527,$ and $\hat{u} = 4.216 + (.5772)(.527) = 4.52.$ Since $\text{Var} (\hat{b}_{n}/b) = .020$ and $\text{Cov} (\hat{a}/b, \hat{b}_{n}/b) = -.004, b* = .527/[1 + .020] = .517, u* = 4.52 - 4.004)(.517) = 4.52, \beta* = 1/.517 = 1.93, \alpha* = \exp(4.52) = 91.8, \text{and } R'(32.46) = .875. For comparison, the MLE’s are $u = 4.53, \hat{b} = .514, \hat{a} = 92.8, \beta = 1.95,$ and $R(32.46) = .878.

3. Distributional Results

An approximate chi square distribution for $2nk_{1,\hat{b}}/b$ was established in [3] for heavily censored samples. In [7] it was shown that asymptotically as $n \rightarrow \infty$ with $h$ fixed, $h\hat{b}/b \sim \chi^2(h)$ where $h = 2/\text{Var}(\hat{b}/b)$. This also has the correct asymptotic distribution as $n \rightarrow \infty$ with $r/n \rightarrow p > 0$, and a simulation study in [7] indicates a good small-sample approximation. An interesting supportive argument is also given in [15, p. 241]. As noted in section 1, the agreement with the RILE is enhanced when $b/\hat{b}$, and the MMSE estimate of $p$ is $(96/100)^{1.93} = 1.89$ and the MMSE estimate of $\beta$ is $(96/100)^{1.93} = 1.85$.

Confidence bounds for other parameters, based upon the simple method, are proposed in [15, p. 245]. In particular, for samples that are not highly censored, a lower confidence bound for $\alpha$ (or $u$) is obtained. The method involves a two-moment chi-square fit to $Q = \exp [(y - u)/b]$ using the mean and variance of the exponential order statistic statistic $\exp [(y - u)/b]$. A confidence bound for $u$, based upon an approximate $F$ variate is derived in [15, p. 249] for highly censored samples. In order to obtain the approximate $F$ variate, a table of variances for the estimator of $u$, obtained by Monte Carlo simulation, is provided in [15, p. 222] for $n = 25/30/60$. The method of section 4 has been used to obtain exact values of $n \text{Var}(\hat{b}/b)$, and these values are provided in Table 5 for $n \equiv 30/60/90$ and $r/n = .1(1.9).$ Exact values for $n = 10, 20$ and $\infty$ are also provided in [8]. After multiplying the simulated values by $n$, the accuracy was found to be comparable to that of the simulated values of $n \text{Var}(\hat{b}/b)$ and $n \text{Cov}(\hat{a}/b, \hat{b}/b)$ as discussed in section 2. It should also be noted that the variances of the estimator of $u$, as proposed in [15] for complete samples, are larger than the corresponding variances for $r/n = .9$. This is evident from the tabulated values in [15, p. 252], and has been further substantiated by calculation of the exact variance for $n = 20$, and the asymptotic variance as $n \rightarrow \infty$ with $s/n \rightarrow .892$. For $n = 20$, the exact variance is .063 for $r/n = 1.0$ and .062 for $r/n = .9$, while the asymptotic variance is $1.290/n$ for $r/n = 1.0$.

| Table 3—Exact values of $n \text{Var}(\hat{b}/b)$ |
|---|---|---|---|---|---|---|---|---|---|
| $r/n$ | .1 | .2 | .3 | .4 | .5 | .6 | .7 | .8 | .9 |
| 30 | 14.62 | 5.931 | 3.452 | 2.436 | 1.851 | 1.471 | 1.204 | 1.008 | .8663 |
| 40 | 13.00 | 5.420 | 3.350 | 2.382 | 1.819 | 1.450 | 1.200 | .9981 | .8541 |
| 50 | 12.18 | 5.269 | 3.290 | 2.350 | 1.800 | 1.437 | 1.218 | .9925 | .8519 |
| 60 | 11.70 | 5.173 | 3.251 | 2.330 | 1.787 | 1.429 | 1.215 | .9888 | .8505 |

TECHNOMETRICS® V. 17, NO. 3, AUGUST 1975
and $1.251/n$ for $r/n = .9$. The value $s$, which was chosen in [7] to minimize the variance of $\hat{b}$ in the complete sample case, is apparently not an optimum choice for the estimator of $\alpha$ as defined in [15]. This suggests that, for larger sample sizes, a more efficient bound or a more powerful test for $\alpha$ might be obtained by censoring 10% rather than using the method proposed in [15] for complete samples.

The following approach can also be used to obtain exact, asymptotically efficient lower confidence bounds for $\alpha$ or $\alpha$, using the estimators $\hat{u}'$ and $\hat{b}'$ from [10]. Confidence bounds for $\alpha$ with $b$ unknown can be based upon values $t = t_s$ such that $\gamma = P[(\hat{u}' - \mu)/\hat{b}' < t] = 1 - P[0 < \hat{u}'/\hat{b} - (\hat{u}' - \mu)/\hat{b}]$. This is equivalent to solving the equation $L(t) = 0$, where $L(t)$ is the function defined in [10] by the equation $\gamma = P[L(t) < \hat{b}'/\hat{u} - (\hat{u}' - \mu)/\hat{b}]$. This can be accomplished, using interpolation, from the tables in [10]. For fixed $\gamma$, simply find the value $t_s$ such that $L(t_s) = \exp(-\exp(0)) = .368$. As a numerical illustration, using the example from section 2, for $r/n = 1$, $u' = 4.52$ and $b' = .523$. For $n = 40$ it is also necessary to interpolate between $n = 30$ and 50 in [10]. For $\gamma = .99$, the interpolated value is $t_{.99} = .40$, and thus $4.52 - (.40)(.523) = 4.31$ is a 99% lower confidence bound for $\alpha$ and $\exp(4.31) = 74.4$ is a 99% lower confidence bound for $\alpha$. For comparison, the corresponding bounds based upon the MLE's are 4.32 and 75.2 respectively. Similar computations for $r/n = .5$ yield the same lower confidence bound 4.21 using either the estimators from [10] or the MLE's.

4. Derivation of Exact Moments

Exact values of $n \text{Var}(\hat{u}'/b)$ and $n \text{Var}(\hat{u}'/b)$ were derived in [7] and [8] respectively for $n = 10$ and 20, using the covariances provided in [13] for $n \leq 25$. These covariances were, in turn, based upon the computational methods of Lieblein [11]. Using these methods, the expected values $E\left(\sum_{i=1}^{r-1} y_i\right)^2$ and $E\left(y_i \sum_{i=1}^{r-1} y_i\right)$ can be obtained by evaluating

$$\frac{\partial^2}{\partial \delta^2} M(s, t)$$

and

$$\frac{\partial^2}{\partial \delta \partial t} M(s, t)$$

respectively for $\theta = t = 0$. These evaluations, of $\sum_{i=1}^{r-1} y_i$ and $y_i$ is derived, and formulas for the moments are obtained by differentiation.

Let $y_1 < y_2 < \ldots < y_r$ denote the smallest $r$ out of $n$ standardized $(u - 0, b - 1)$ extreme-value order statistics. Then the joint moment generating function of $\sum_{i=1}^{r-1} y_i$ and $y_r$ is

$$M(s, t) = E\left[\exp\left(s \sum_{i=1}^{r-1} y_i + ty_r\right)\right].$$

After the substitution $x_i = \exp(y_i)$ for $i = 1, 2, \ldots, r$ we have

$$M(s, t) = \frac{1}{n!/(n - r)!} \int_0^\infty x^r \exp[-(n - r + 1)x] G_{r-1}(x; s) \, dx,$$

where

$$G_{r-1}(x; s) = \int_0^x \cdots \int_0^{x_r} \prod_{i=1}^{r-1} [x_i \exp(-x_i)] \, dx_1 \cdots dx_{r-1}.$$
followed by the substitution \( u = 1 - \exp (-x) \),
yield the following formulas:

\[
E\left( \sum_{i=1}^{r-1} y_i \right)^2 = K_3 \int_0^1 u^{-r}(1 - u)^{-r} |I_1(\lambda_0)|^2 \, du \\
+ K_2 \int_0^1 u^{-r}(1 - u)^{-r} |I_2(\lambda_0)| \, du,
\]

and

\[
E\left( \sum_{i=1}^{r-1} y_i \right) = K_2 \int_0^1 u^{-r}(1 - u)^{-r} \ln (\lambda_0) I_1(\lambda_0) \, du,
\]

where \( \lambda_0 = -\ln (1 - u) \),

\[
I_1(x) = \int_0^x \ln t \exp (-t) \, dt,
\]

\[
I_2(x) = \int_0^x (\ln t)^2 \exp (-t) \, dt,
\]

and \( K_i = \frac{n!}{(n - r)!} \cdot \frac{(r - j)!}{(r - j)!} \) for \( j = 2, 3 \). The functions \( I_1(\lambda_0) \) and \( I_2(\lambda_0) \) can be evaluated as follows: \( I_1(\lambda_0) = u \ln (\lambda_0) + A(u) \) and \( I_2(\lambda_0) = u(\ln (\lambda_0))^2 + 2 \ln (\lambda_0) A(u) + 2B(u) \) where

\[
A(u) = \sum_{i=1}^n (-1)^i \lambda_0^{i+1}/i!i!,
\]

and

\[
B(u) = \sum_{i=1}^n (-1)^{i+1} \lambda_0^{i+1}/i!i^2.
\]

From these formulas, the desired expected values can be obtained by numerical integration. The expected values \( E(y_i) \), \( E(y_i^2) \), and \( E(\sum_{i=1}^{r-1} y_i) \) can also be obtained by this method, although it is simpler to calculate these from values in [16].

Exact values of \( n \text{ Var } (\hat{b}/b) \), \( n \text{ Cov } (\hat{a}/b, \hat{b}/b) \), and \( n \text{ Var } (\hat{a}/b) \) are provided in Tables III, IV, and V respectively for \( n = 30(10)60 \) and \( r/n = .1(.1).9 \).

REFERENCES


